Combinatorial Optimization in Computer Vision
Lecture: F. R. Schmidt and C. Domokos
Exercises: T. Möllenhoff and T. Windheuser
Winter Semester 2015/2016

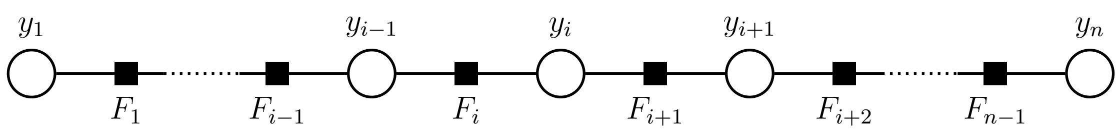
Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 5

Room: 02.09.023
Tuesday, 24.11.2015, 14:15-15:45
Submission deadline: Tuesday, 24.11.2015, 11:15 , Room 02.09.023

Probabilistic Inference (5 Points)

Exercise 1 (Inference on chains, 2 Points)**.** Consider the following factor graph, which is a chain:



The joint distribution can be written in the form

$$p(\mathbf{y}) = \frac{1}{Z} F_1(y_1, y_2) F_2(y_2, y_3) \cdot \ldots \cdot F_{n-1}(y_{n-1}, y_n),$$

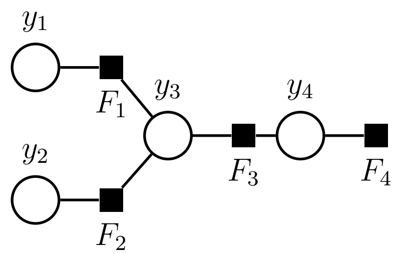
where $Z = \sum_{\mathbf{y}} \prod_{i=1}^{n-1} F_i(y_i, y_{i+1})$ denotes the partition function. Show that the marginal distribution $p(y_i)$ decomposes into the product of two factors:

$$p(y_i) = \frac{1}{Z} r_{F_i \to y_i}(y_i) r_{F_{i+1} \to y_i}(y_i).$$

Solution. Using the definition of $p(y_i)$ we have:

$$\begin{aligned}
p(y_i) &= \sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} p(\mathbf{y}) \\
&= \sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} \frac{1}{Z} F_1(y_1, y_2) F_2(y_2, y_3) \cdot \ldots \cdot F_{n-1}(y_{n-1}, y_n) \\
&= \frac{1}{Z} \sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} F_1(y_1, y_2) F_2(y_2, y_3) \cdot \ldots \cdot F_{n-1}(y_{n-1}, y_n) \\
&= \frac{1}{Z} \left[\sum_{y_{i-1}} F_{i-1}(y_{i-1}, y_i) \ldots \left[\sum_{y_2} F_1(y_2, y_3) \left[\sum_{y_1} F_1(y_1, y_2) \right] \right] \ldots \right] \cdot \qquad (1) \\
&\quad \left[\sum_{y_{i+1}} F_i(y_i, y_{i+1}) \ldots \left[\sum_{y_n} F_{n-1}(y_{n-1}, y_n) \right] \ldots \right] \\
&= \frac{1}{Z} r_{F_i \to y_i}(y_i) r_{F_{i+1} \to y_i}(y_i).
\end{aligned}$$

Exercise 2 (Max-sum algorithm, 3 Points)**.** Execute the max-sum algorithm to find a maximizing configuration $\mathbf{y} \in \{0, 1, 2\}^4$ of the following factor graph:



The factors are defined through the following:

$$\begin{aligned} &F_x(y_i, y_j) = \exp\left(-(|y_i - y_j| + (c_x - y_i)^2)\right), x \in \{1, 2, 3\}, \\ &c_1 = 0, c_2 = 1, c_3 = 1, c_4 = 2, \\ &F_4(y_4) = \exp(-(c_4 - y_4)^2). \end{aligned} \tag{2}$$

Show the intermediate steps in detail. Pick y_4 as the root node.

Solution. We consider y_4 as the root node. Starting with the leaf nodes, we then have the following messages:

$$r_{F_4 \to y_4}(y_4) = \log F_4(y_4) = -(c_4 - y_4)^2, \tag{3}$$

$$q_{y_1 \to F_1}(y_1) = 0, \tag{4}$$

$$r_{F_1 \to y_3}(y_3) = \max_{y_1} \{\log F_1(y_1, y_3) + q_{y_1 \to F_1}(y_1)\}, \tag{5}$$

$$q_{y_2 \to F_2}(y_2) = 0, \tag{6}$$

$$r_{F_2 \to y_3}(y_3) = \max_{y_2} \{\log F_2(y_2, y_3) + q_{y_2 \to F_2}(y_2)\}, \tag{7}$$

$$q_{y_3 \to F_3}(y_3) = r_{F_1 \to y_3}(y_3) + r_{F_2 \to y_3}(y_3), \tag{8}$$

$$r_{F_3 \to y_4}(y_4) = \max_{y_3} \{\log F_3(y_3, y_4) + q_{y_3 \to F_3}(y_3)\}. \tag{9}$$

A quick calculation shows that the messages are given as:

	0	1	2
$r_{F_4 \to y_4}$	-4	-1	0
$r_{F_1 \to y_3}$	0	-1	-2
$r_{F_2 \to y_3}$	-1	0	-1
$q_{y_3 \to F_3}$	-1	-1	-3
$r_{F_3 \to y_4}$	-2	-1	-4

Hence, the maximizing energy is given as:

$$E(\hat{\mathbf{y}}) = \max_{y_4} \{r_{F_3 \to y_4}(y_4) + r_{F_4 \to y_4}(y_4)\} = -2. \tag{10}$$

To find a maximizing configuration, we have the following sequence of updates:

$$\widehat{y}_4 = \underset{y_4}{\operatorname{argmax}} \{r_{F_3 \to y_4}(y_4) + r_{F_4 \to y_4}(y_4)\} = 1. \tag{11}$$

$$\widehat{y}_3 = \underset{y_3}{\operatorname{argmax}} \{r_{F_2 \to y_3}(y_3) + r_{F_1 \to y_3}(y_3) + \log F_3(y_3, 1)\} = 1. \tag{12}$$

$$\widehat{y}_2 = \underset{y_2}{\operatorname{argmax}} \{\log F_2(y_2, 1)\} = 1, \tag{13}$$

$$\widehat{y}_1 \in \underset{y_1}{\operatorname{argmax}} \{\log F_1(y_1, 1)\} = \{0, 1\}. \tag{14}$$

Thus, two global maximizers are given by:

$$\widehat{\mathbf{y}} \in \{(0, 1, 1, 1), (1, 1, 1, 1)\}. \tag{15}$$

Roof Duality (5 Points)

Exercise 3 (Roof duality, 5 Points)**.** Consider the following pseudo-Boolean energy function $f : \mathbb{B}^5 \to \mathbb{R}$:

$$f(x_1, \dots, x_5) = 10 - 4x_1 - 4x_3 - 2x_4 + 4x_1x_2 - 2x_2x_3 + 4x_3x_4 - 2x_4x_5.$$

a) Show that f is not submodular.

b) Find the global minimizer $\widehat{x} \in \mathbb{B}^5$ of f using roof duality.

c) Show that f is submodular with respect to x_1, $\bar{x}_2$, $\bar{x}_3$, x_4, x_5.

Solution.

1. Among others, the second derivative $\frac{\partial^2 f}{\partial x_1 x_2} = 4$ is positive. Hence f is not submodular.

2. First we rewrite the energy as a posiform:

$$\begin{aligned}
&f(x_1, \dots, x_5) = 10 - 4x_1 - 4x_3 - 2x_4 + 4x_1x_2 - 2x_2x_3 + 4x_3x_4 - 2x_4x_5 \\
&= 10 - 4(1 - \bar{x}_1) - 4(1 - \bar{x}_3) - 2(1 - \bar{x}_4) + 4x_1x_2 - 2(1 - \bar{x}_2)x_3 + 4x_3x_4 - 2(1 - \bar{x}_4)x_5 \\
&= 4\bar{x}_1 + 4\bar{x}_3 + 2\bar{x}_4 + 4x_1x_2 - 2(1 - \bar{x}_3) + 2\bar{x}_2x_3 + 4x_3x_4 - 2(1 - \bar{x}_5) + 2\bar{x}_4x_5 \\
&= -4 + 4\bar{x}_1 + 6\bar{x}_3 + 2\bar{x}_4 + 2\bar{x}_5 + 4x_1x_2 + 2\bar{x}_2x_3 + 4x_3x_4 + 2\bar{x}_4x_5.
\end{aligned}$$

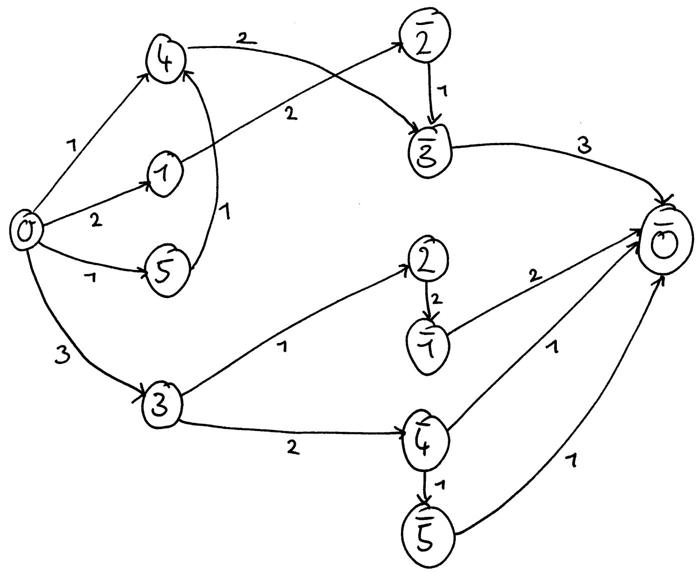
   This can be written with $V = \{0, \bar{0}, 1, \bar{1}, \dots, 5, \bar{5}\}$ as:

$$\Phi(x) = C_0 + \sum_{i,j \in V} C_{ij} x_i x_j, \tag{16}$$

   with

$$\begin{aligned}
&C_0 = -4, C_{0\bar{1}} = 4, C_{0\bar{3}} = 6, C_{0\bar{4}} = 2, C_{0\bar{5}} = 2, \\
&C_{12} = 4, C_{\bar{2}3} = 2, C_{34} = 4, C_{\bar{4}5} = 2.
\end{aligned}$$

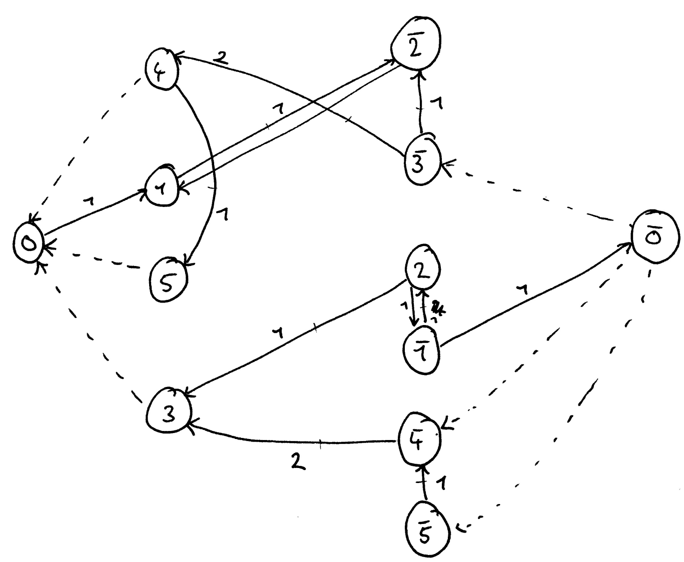
   The network associated with (16) is given as the following:



We have the augmenting paths (each with unit flow):

$$
\begin{aligned}
&0 \to 4 \to \bar{3} \to \bar{0}\\
&0 \to 1 \to \bar{2} \to \bar{3} \to \bar{0}\\
&0 \to 5 \to 4 \to \bar{3} \to \bar{0}\\
&0 \to 3 \to \bar{4} \to \bar{0}\\
&0 \to 3 \to \bar{4} \to \bar{5} \to \bar{0}\\
&0 \to 3 \to 2 \to \bar{1} \to \bar{0}.
\end{aligned}
$$

This yields the following residual network:



Since x_1 and $\bar{x}_2$ are connected to the source we have $x_1 = 1$ and $x_2 = 0$. Furthermore the residual network corresponds to the posiform

$$\Phi'(x) = 2\bar{x}_1 + 2x_1x_2 + 2\bar{x}_1\bar{x}_2 + 2x_2\bar{x}_3 + 4\bar{x}_3\bar{x}_4 + 2x_4\bar{x}_5$$

Substituting the above into the posiform finally yields:

$$4\bar{x}_3\bar{x}_4 + 2x_4\bar{x}_5$$

It can be seen that a minimizing configuration is given as $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 1$, $x_5 = 1$, giving an energy of 2 when substituted in the original energy (16).

3. To check submodularity, it is enough to consider the pairwise terms:

$$\begin{aligned} f_{\text{pairwise}}(x_1, \ldots, x_5) &= 4x_1x_2 - 2x_2x_3 + 4x_3x_4 - 2x_4x_5 \\ &= 4x_1(1-\bar{x}_2) - 2(1-\bar{x}_2)(1-\bar{x}_3) + 4(1-\bar{x}_3)x_4 - 2x_4x_5 \\ &= 4x_1 - 4x_1\bar{x}_2 - 2(1 - \bar{x}_3 - \bar{x}_2 + \bar{x}_2\bar{x}_3) + 4x_4 - 4\bar{x}_3x_4 - 2x_4x_5 \\ &= \underbrace{\ldots}_{\text{unary terms}} - 4x_1\bar{x}_2 - 2\bar{x}_2\bar{x}_3 - 4\bar{x}_3x_4 - 2x_4x_5. \end{aligned}$$

It can be seen that all second derivatives are negative, thus f is submodular with respect to x_1, $\bar{x}_2$, $\bar{x}_3$, x_4, x_5.

Programming (2 weeks time, 10 Points)

Presentation of the programming exercise will be on Tuesday, December 1st.

Exercise 4 (Texture Denoising, 10 Points)**.**
Let $\mathcal{I} : \Omega \to [0, 1]$, where $\Omega = \{0, \ldots, W-1\} \times \{0, \ldots, H-1\}, W, H \in \mathbb{N}$, be a binary image with additive Gaussian noise. We would like to recover the original binary image $\mathcal{I}_0$ by removing the noise.

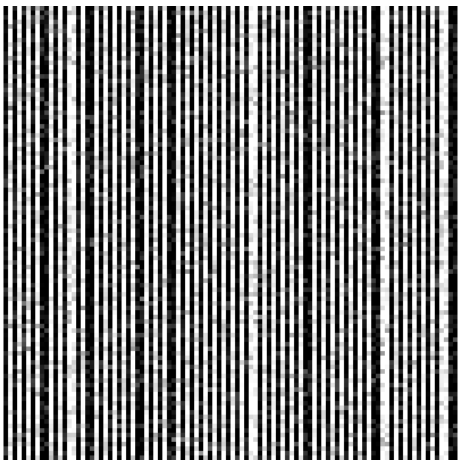
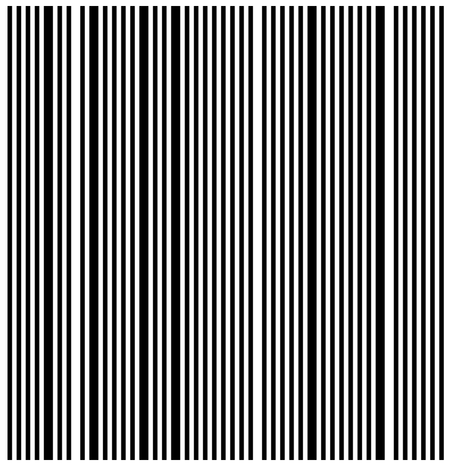


image $\mathcal{I}$ with noise


original image $\mathcal{I}_0$

Fortunately, we know that the original image contains mostly vertical stripes and only black or white pixels, i.e. $\mathcal{I}_0 : \Omega \to \mathbb{B}$. Therefore, we can describe the denoised image as the optimizer of the following energy function:

$$E(X) = \lambda \sum_{i,j} f_{i,j}(X(i,j)) + \sum_{i,j} c_{\mathrm{h}}(X(i,j), X(i+1,j)) + c_{\mathrm{v}}(X(i,j), X(i,j+1)), \tag{17}$$

where $X : \Omega \to \mathbb{B}$ is a binary image and $\lambda > 0$ is a positive scalar parameter. The functions $f_{i,j} : \mathbb{B} \to \mathbb{R}$ encode the data given by image $\mathcal{I}$ and are defined by

$$f_{i,j}(x) = (\mathcal{I}(i,j) - x)^2. \tag{18}$$

The pairwise regularizers $c_{\mathrm{h}}, c_{\mathrm{v}} : \mathbb{B} \times \mathbb{B} \to \mathbb{R}$ encode the knowledge about the stripe pattern of the original image. The horizontal term c_{h} prefers neighboring pixels that have different intensities:

$$c_{\mathrm{h}}(x_1, x_2) = 1 - |x_1 - x_2|.$$

The vertical term c_{v} prefers neighboring pixels that have the same intensities:

$$c_{\mathrm{v}}(x_1, x_2) = |x_1 - x_2|.$$

Which energy is submodular? c_{h} or c_{v}?

Write a program that tries to find the optimizer of energy function (17). Since half of the pairwise terms are non-submodular try the following strategies:

a) Remove the non-submodular terms from energy function (17) and use the basic graph-cut algorithm.

b) Use QPBO to improve the result of a).

You can use the code and the images inside `05_supp.zip` from the lecture website as a starting point. The file `texturedenoising.cpp` already includes the code to load the images and textures and to compute the pairwise energy functions.

The exercise uses a simplified version of the model in Cremers and Grady [1] `https://vision.in.tum.de/_media/spezial/bib/cremers_grady_eccv06.pdf`.

References

[1] Cremers, D., Grady, L.: Statistical priors for combinatorial optimization: efficient solutions via Graph Cuts. In Leonardis, A., Bischof, H., Pinz, A., eds.: European Conference on Computer Vision (ECCV). Volume 3953 of LNCS., Graz, Austria, Springer (2006) 263–274